DOCUMENT RESUME

ED 258 448 FL 015 056

AUTHOR Lombardo, Marie
TITLE The Effectiveness of an Informal Reading Inventory in Identifying the Functional Reading Levels of Bilingual Students. Bilingual Education Paper Series Vol. 2 No. 10.
INSTITUTION California State Univ., Los Angeles. Evaluation, Dissemination and Assessment Center.
SPONS AGENCY Office of Bilingual Education and Minority Languages Affairs (ED), Washington, D.C.
PUB DATE May 79
NOTE 62p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *Bilingual Students; Cloze Procedure; Comparative Analysis; *Informal Reading Inventories; *Reading Achievement; Reading Comprehension; *Reading Tests; Secondary Education; Secondary School Students; Spanish Speaking; Standardized Tests; Test Construction; *Test Reliability; Test Validity

ABSTRACT

A study was undertaken to (1) examine the development and construction of a Group Informal Reading Inventory to predict the reading comprehension levels (independent, instructional, and frustration) of junior high school bilingual students for the purpose of reading instruction; and (2) validate the inventory through a three-way correlational study comparing the comprehension results with those of a cloze test, a standardized test, and a questionnaire by which teachers estimate students' reading levels. The study involved 50 bilingual students of predominantly English- and Spanish-speaking, low- and middle-income backgrounds in an urban school. All had been instructed in Spanish until they gained English language proficiency; then they were mainstreamed into the English curriculum. It was discovered that the students were all functioning far below their developmental grade levels and their assigned present grade levels, and native language grades were lower than those in English. It is recommended that (1) a decision be made for each individual student as to whether he should be taught in two languages or, if his native language skills are insufficient to transfer to English as a second language, whether he should be taught in English; (2) testing for reading and content areas be administered regularly to monitor progress; and (3) there be careful regulation of the timing, techniques, content, materials, and evaluation of bilingual instruction. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ISSN 0161-3707

Vol. 2 No. 10
May 1979

ED258448

Bilingual Education PAPER SERIES

THE EFFECTIVENESS OF AN INFORMAL READING INVENTORY IN IDENTIFYING THE FUNCTIONAL READING LEVELS OF BILINGUAL STUDENTS

Marie Lombardo
Boston University

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☒ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Evaluation, Dissemination and Assessment Center
California State University, Los Angeles
5151 State University Drive
Los Angeles, CA 90032

FL 015 056

The subject of this publication was supported in whole or in part by the Office of Education, United States Department of Health, Education, and Welfare. The opinions expressed herein do not necessarily reflect the position or policy of the Office of Education; no official endorsement by the Office of Education should be inferred.

This publication was printed with funds provided by the Bilingual Education Act, Title VII of the Elementary and Secondary Education Act of 1965, as amended by Public Law 93.380.

THE EFFECTIVENESS OF AN INFORMAL READING INVENTORY IN IDENTIFYING THE FUNCTIONAL READING LEVELS OF BILINGUAL STUDENTS*

Marie Lombardo

INTRODUCTION

The main purpose of this study is twofold:

1. The development and construction of an original Group Informal Reading Inventory (GIRI) in predicting the comprehension reading levels (Independent, Instructional, and Frustration) of junior high bilingual students for the purpose of reading instruction.
2. The validation of the GIRI through a three-way correlational study that will compare the comprehension results of the GIRI with those of a *Cloze Test*, a standardized test (the *Stanford Diagnostic Reading Test*, 1976), and a questionnaire in which the teachers estimate students' reading levels.

(For the data on validation of the GIRI, the reader may consult the original dissertation study.)

---

*This paper is based on the author's doctoral dissertation entitled, "The Construction and Validation of the English as a Second Language Assessment Battery: The Receptive Area," Boston University, 1979.

Justification

In recent years, it has been reported that in spite of the fact that bilingual programs do exist for bilingual students, these students are still performing below grade level (Herbert, 1977) and, consequently, many become frustrated and eventually drop out of school (United States Commission on Civil Rights, 1971). The frustration that these students encounter is based upon the fact that: (a) their reading levels are not properly assessed, their individual reading needs are not met, and as a result they are presented with materials that are too difficult for them; and (b) students are often grouped for reading according to standardized test results. These results are not accurate because standardized measures tend to overestimate the reading levels of students, thus placing these students at their frustration rather than their instructional reading level (Wiechelman, 1971; Motta *et al.*, 1974). Also, standardized measures may not be appropriate if they were not designed for the bilingual students to be assessed:

> When standardized measures are used, results may be reasonably reliable and valid, but interpretation of the individual student's performance may not be possible if people like the student were not part of the group on which the test was normed. Using a standardized test to assess reading ability for instance, is not appropriate when the student is not a native speaker of English. While the test might be used to compare the performance of a foreign-born bilingual with American monolinguals for diagnostic purpose, the resulting score is not a meaningful estimate of reading ability of the bilingual student. (Morishima and Mizokawa, 1977, p. 2)

Condon (1975) adamantly states that little research has been conducted in the United States relating to the influence of cultural factors on the fairness of the assessment of culturally different students. She purports that most of the problems raised in assessment have come from bilingual or compensatory programs, which require the use of standardized tests to measure the functional progress of bilingual students. According to Condon (1975), the problem is the contextual misconstruction of items; there is a neglect of culture, misinformation, an erroneous presentation of data, and finally bias, or the presentation of distorted impressions of the foreign culture. As a result, bilingual students are unfairly assessed in terms of the content of these measures and the uses to which their results are applied.

Surely, the case for reading assessment instruments to be specifically designed for bilingual students is apparent. However, to place the need for such instruments in proper perspective, the evolution of this problem in bilingual education must be presented in an historical overview.

Historically, the United States has had many changes in its language policies. Until 1880, there was no language policy enforcing the sole use of English. There was tolerance for the pupils' primary languages (Spanish, German, Portuguese, French, Dutch, and Basque). (Anderson and Boyer, 1970)

At about 1917, with the occurrence of World War I, other languages were considered a threat to nationalism. As a result of the war, many people from southern Europe immigrated

to the United States. During this period, states such as Connecticut and Massachusetts began enforcing the "use of English" policies and the requirement of literacy before voting (passed in 1855 and 1857, respectively). But the English-only policy, also known as the "melting pot idea," became intensified during World War I. This concept was not altered until 1954 with the *Brown* v. *Board of Education* Civil Rights decision, which called for equal educational opportunities for all races. The most important impact of the case was the fact that it agitated minority groups to also demand social and political opportunities. Through the 1960's evidence (Kobrick, 1972; *The Way We Go to School*, 1970) collected indicated that the traditional English monolingual educational system did not meet the needs of over 2.5 million students of language backgrounds other than English.

In 1968, the Bilingual Education Act was passed as part of ESEA. This act, later revised in 1974, stated that the educational program allow non-English-speaking children to be taught in English and the native language so as to facilitate their progression through school. Language is the key word in this act. The factor of language was further considered in court cases, whereby parents wanted to ensure that their children were taught the native lanuage before grade eight. However, it was not until the *Lau* v. *Nichols* (1974) case that guidelines for bilingual education were established. This case was filed in San Francisco on behalf of the 1,800 Chinese students who were not receiving appropriate language

assistance to enable them to compete on equal grounds with their English-speaking classmates. The Department of Health, Education, and Welfare declared that the recipients of federal financial aid could not "restrict an individual in any way in the enjoyment of any advantage or privilege enjoyed by others receiving any services, financial aid, or any other benefit under the program" (*Lau v. Nichols*, 1974, p. 566).

To comply with the decision of the Court, the San Francisco Unified School District, with a citizen's task force, designed guidelines for school districts to follow in the case of students "whose home language is other than English" (Office of Civil Rights, 1975). The guidelines known as the "Lau Remedies" included the identification of the student's:

1. Language dominance according to five categories:
    a. Monolingual speaker of the language other than English (speaks the language other than English exclusively).
    b. Predominantly speaks the language other than English (speaks mostly the language other than English, but speaks some English).
    c. Bilingual (speaks both the language other than English and English with equal ease).
    d. Predominantly speaks English (speaks mostly English, but some of the language other than English).
    e. Monolingual speaker of English (speaks English exclusively). (Lau Remedies, Office of Civil Rights, 1975, p. 2)
2. Frequency of use.
3. Diagnosis and prescriptive approach for proposing an educational program:
    a. At the elementary and intermediate levels the program may be transitional, bilingual/bicultural, or multilingual/multicultural.

b. At the secondary level the program may be bilingual, transitional, ESL or any of the above mentioned combinations.

4. An outline of requirements of personnel teaching.

5. A system for notifying the parents of the student's program.

6. A way to evaluate the program. (Lau Remedies, Office of Civil Rights, 1975, pp. 2-3)

Although these remedies were certainly a step in the right direction, the problem is that once students are selected according to the remedies (categories a, b, or c) and placed in a program, it is not required that their language competence in language (listening, speaking, reading, and writing) be assessed. However, from an educator's standpoint, it is clear that assessment should be required. More specifically, there is a pressing need for reading instruments to be developed and validated especially for bilingual students taking into account their interests and backgrounds in the assessment of their functional reading levels (levels at which each of the students can function adequately in the classroom). As will be demonstrated, informal reading inventories (IRI) are proposed to meet such a need.

The need for IRI[illegible]n assessing students' functional reading levels for the purpose of grouping and of matching materials to students' needs in monclingual and bilingual classrooms has been the most urgent request of reading authorities. Harris (1961) claimed that the use of inappropriate reading materials was the most frequent cause of reading difficulties faced by experienced teachers. Dechant (1970) declared that

10-15 million children in regular classrooms attempted to read books that were too hard for them; and that the use of inappropriate textbooks interfered with reading progress.

The claim that standardized tests could give the instructional levels for individual children was disputed (Smith, 1961; Tucker, 1975). If teachers did careful readability studies of the materials to be used in the classroom and devised IRI's and administered these, they would discover the instructional reading levels of their students. They could then match materials to student levels and needs, thus making instruction more effective.

In an attempt to comply to the need for reading instruments, this study will attempt to propose and validate an original Group Informal Reading Inventory (GIRI) for bilingual students.

REVIEW OF THE LITERATURE

In reviewing the existing literature, it is apparent that IRI's constitute a controversial issue. Researchers have focused on different aspects of IRI's and have formulated conflicting opinions on the definition, construction, administration, and scoring. But, in spite of these disputed areas, researchers are in unison regarding its purpose, need, and validity. For the purpose of this study, the areas of conflict will first be described followed by the purpose, need, and the validity of IRI's, when they are compared with the *Cloze* and standardized tests.

In identifying the first conflict area--that of defining the IRI--it is agreed by most reading authorities that IRI's are informal diagnostic measures because usually they are neither normed nor standardized (Johnson and Kress, 1965). More specificially, Johnson and Kress describe an IRI as: (a) *informal* because norms have not been established and the performance of one student is not judged against that of others, but against some standard of mastery; (b) *reading* because it evaluates the student's ability to manipulate ideas represented by words in the receptive and expressive language areas; and (c) *inventory* because it reports the students' complete comprehensive performance in reading, language, and thinking skills.

Although an IRI is defined as a non-standardized reading test through which a student's reading performance is evaluated against predetermined standards (McCracken, 1967), researchers (McCracken, 1964, 1970; Botel, 1961; Silvaroli, 1965) have developed and standardized their own IRI's by establishing the reliability and validity of their tests through various studies.

McCracken (1964) conducted a study to validate his "Standard Reading Inventory," which is based on the vocabulary of three basal readers, and tested the validity of its passages with well-known readability formulae. He then improved the content validity of the inventory by controlling the vocabulary, sentence length, content, and style of the reading selections and obtained norming data by administering these oral reading selections to 664 students in grades 1-6. The significant differences found in student performance as para-

graphs increased in difficulty was quite substantial. Since there were alternate forms of the inventory, reliability was obtained by having two examiners administer each form to 60 elementary school children. Reliability correlations between the two forms (for the Independent, Instructional, and Frustration Reading Levels) ranged from .86 to .91. The correlations between the two forms for the eight reading sub-skills measured by the inventory ranged from .68 for Word Recognition errors to .99 for Vocabulary in isolation. It was determined by the results of this study that the "Stanford Reading Inventory" reliably estimates students' functional reading levels.

Botel (1961) cross-validated his reading test and readability measures through the use of correlational and matching procedures. The correlations among all the tests and readability measures were unusually high for both McCracken's and Botel's tests. In a later study, McCracken and Mullen (1970), further validated the IRI's by compiling the data from the administration of the "Standard Reading Inventory" (SRI), the "Botel Inventory," and the *Stanford Achievement Tests* to 171 male and female students from grades 1-6. Results indicated that there was concurrent validity between the mean levels of the *Stanford Achievement Tests* and the instructional levels obtained from the SRI and the Botel Inventory. The data also confirmed that the instructional reading levels can be measured validly.

Most researchers, however, have not advocated standardization of IRI's. In addition to the explanation of IRI, it

is necessary to delineate its purpose, which is three-fold: (a) to assess the student's functional level and the student's strengths and weaknesses for purposes of instruction (Johnson and Kress, 1965); (b) to estimate the independent, instructional, and frustration reading level for each student so that reading materials can be matched accordingly (Betts, 1957; Botel, 1969; Beldin, 1970; Dipay, 1964; Kender, 1968); and (c) to serve as a placement instrument in grouping students according to their appropriate reading levels (Walter, 1974; Betts, 1940; McCracken, 1967). Both Walter (1974) and Pilulski (1974) claim that there is no need to standardize IRI's since there is face validity in IRI's when their passages and questions are sampled from the materials used, or be to used, in the classroom.

In addressing the second area of conflict--construction of IRI's--reading authorities have been divided in their opinions; some have utilized and advocated the use of classroom materials, others have created their own reading selections for IRI's. Motta *et al.* (1974) advocate that in preparing IRI's, the materials to be used in the curriculum should be utilized. They warn that in selecting these materials, one should take into account the elements of interest, culture, and language structure in dealing with non-native speakers. They especially emphasize that an analysis of semantics, lexicons, and syntactical structure should be conducted in order to ensure the appropriateness of these materials for non-native speakers.

On the other hand, researchers--McCracken (1964), Botel (1961), and O'Brien (1970)--have reviewed basal readers and

found it most feasible to create original IRI's based on their subjects' interests. Otto *et al.*(1973), also agreed that an IRI does not have to be constructed from the basal reader series for students of the upper elementary and secondary levels.

O'Brien (1970) conducted a study to confirm that her original IRI's were more effective than those whose passages were directly extracted from classroom basal readers. She studied traditional IRI's and suggested a new method for devising them in which words are taken from basal readers and incorporated into an original paragraph. To discover the usefulness of this procedure, a traditional IRI and one based on the new method were administered to a group of second and fourth graders. The results indicated that the new IRI's: (a) presented fewer words per selection, (b) presented more new words in each selection, (c) required fewer selections to be read, and (d) gave an instructional level score (in ten cases) lower than did the traditional IRI.

Once the context for the IRI has been selected, the next consideration is the type of IRI to be constructed. Types vary from those based on graded word lists (Silvaroli, 1965; Motta *et al.*, 1974) to a series of passages and corresponding comprehension questions (McCracken, 1964; Botel, 1961, Silvaroli, 1965). Based on research and sensitivity to bilingual students' needs, the researcher decided to create original stories for the purpose of the study. The advantages of the stories are:

1. Context was provided to facilitate reading. Thonis (1976) contends that it is simpler for speakers of other languages to read vocabulary that represents content.
2. Original themes were employed to motivate and interest the non-native speakers.
3. Lexicon (vocabulary), semantics (meaning), and syntax (grammatical structures) were manipulated and adjusted for these students.
4. Dialogue was used as context in the stories so that students could be presented with a natural transition from oral to written language (Ruddell, 1965).

When one wishes to construct the story type of IRI, Otto *et al.* (1973) recommend the following passage lengths: (a) passages of 30-100 words followed by five comprehension questions for the primer level, (b) passages of 150 words followed by six comprehension questions for second and third grade levels, and (c) passages of 150 words followed by eight comprehension questions for levels above fourth grade. The reason that passages are increased in length are that students of limited-English-speaking ability should not be frustrated by long selection. However, longer passages are needed to facilitate the construction of higher level questions. For the purpose of this study, the story type of IRI was constructed. The number of words for the passages and the number of comprehension questions were increased at each level because the GIRI was intended for secondary school students.

After the passages have been constructed, readability indices are computed for each passage. Measures of readability range from the Spache (1953), which is intended for the third grade level and below; the Dale and Chall (1948), which presupposes a reading level of fourth grade; the Flesch (1948), intended for grades 4-16; the Fry (1968) for grades 1-13; and the Smog (cited in Vaughan, 1976) for grades one through college. Pauk (cited in Vaughan, 1976) tested the Smog, Dale-Chall, and Fry readabilities on the same passages of 20 articles and found that: (a) the Dale-Chall and Fry scores were in agreement, (b) the Smog scores disagreed with the Fry and Dale-Chall, and (c) the Smog scores tended to be two grade levels above the Dale-Chall and Fry in estimating the grade levels of the passages. Based on these conclusions, the researcher decided to use the Fry for the primer to fourth grade passages, and the Dale-Chall for passages above the fourth grade level. The rationale for including passages as low as the primer level was to help the bilingual student meet with initial success. Thonis (1976) indicated that pupils who have limited-English-speaking abilities need materials with a lower readability rating (about two to three years) below that of English speakers.

Once the grade level of each passage has been determined, comprehension questions are prepared. The questions deal with the cognitive and affective domains so that the students' knowledge and personalized perspective based on their individual experiences is examined. For each passage: (a) questions

dealing with the cognitive domain were hierarchically arranged from literal to synthesis levels of comprehension, according to Bloom (1956) and Barret's (cited in Lapp and Ramsey, 1976) taxonomies; and (b) questions dealing with the affective domain ranged from simple awareness of situations to the complex determination of one's *own* values and philosophy of life, according to Krathwohl's *et al.* taxonomy (1964).

The questions were constructed according to Lindvall's (1967) model of Bloom's Taxonomy, which delineates the testing objectives and the most expedient way to elicit responses. The model outlines and examines the levels as follows:

1. The knowledge level deals with the student's ability to recall terms, facts, rules, principles, and other generalizations. The testing objectives are for the student to be able to name, list, state, describe, and define material presented.

2. The comprehension level is concerned with the pupil's ability to understand a given content, put it in the pupil's own words, summarize it, and explain it. The testing objectives are for the student to be able to translate, give examples, illustrate, interpret, summarize, and explain given materials.

3. At the application level, the focus is on the student's ability to use rules, methods, procedures, principles, and other types of generalization to produce or explain given consequences or to predict the results of a given situation. The testing objectives are to be able to solve, predict, develop, explain, and apply knowledge in a given situation.

4. At the synthesis level, the pupil's ability to develop, create, or produce something is tested. Specific testing objectives are to develop a plan, write a paper, produce, create, or demonstrate a particular work (Lindvall, 1967, pp. 19-20).

For the intensive purpose of this study and fulfillment of the testing objectives at each level of cognition, multiple-choice items were constructed for the following reasons: (a) to provide a stimulus in order to evoke recall, (b) to facilitate testing since it is easier for non-native speakers of English to recognize and identify rather than retrieve information, (c) to evaluate a greater variety of abilities, (d) to reduce guessing by providing several alternatives (Lindvall, 1967), and (e) to provide an objective method of scoring, a notion supported by Lowell (1969) who advocated that more attention should be given to identifying reading performance in ways that do not rely upon the examiner's judgement.

Once the IRI has been constructed, the third area of conflict to be examined is that of administering the IRI. The issues involved are: (a) the technique to be adopted--whether students are to read the passages orally, silently, or listen to the teacher read then and then respond to comprehension questions; (b) the procedure for administration--whether the test is to be individual or in group; and (c) the choice of the examiner to administer IRI.

In addressing the problem of the technique for administering the IRI, it seems that researchers (Burmeister, 1974; Lowell, 1969; Dunkeld, 1970) were in unison that the value of employing

oral IRI technique above sixth grade is questionable. Dunkeld (1970) strongly contended that comprehension rather than word recognition scores obtained from oral reading seemed to be more valid indicators of passage difficulty above the fifth grade level. Furthermore, the technique of reading IRI's orally has not been supported in literature for several reasons. There is a lack of uniformity among examiners as to whether passages should be read aloud or silently first; the problem is obvious, students reading the passages silently first would, of course, perform better. Also, the true word recognition score is not accurately identified after the student has read several passages orally. Dunkeld (1970) conducted a study and found that with the practice of reading several passages, student's word recognition scores increased. The problem is evident here; one examiner could choose the initial word recognition scores as indicators of the student's reading levels and another may assume the latter scores to be indicators of the student's reading levels.

Along with the method of administering the oral IRI, the major controversy centers around the criteria for scoring oral reading errors. There is agreement on the behavioral characteristics to be considered, but there is disagreement as to the definition of an oral reading error (Wilson, 1972).

The behavioral characteristics are listed as mispronunciations, substitutions, omissions, insertions, regressions, hesitations, and punctuation (Johnson and Kress, 1965; Burmeister, 1974; McCracken, 1967). However, the meaning of these terms

varies from one examiner to the other. Because of this lack of agreement, both early and contemporary reading experts have placed less significance on reading errors as determiners of the students' reading levels.

Early researcher (Betts, 1936) believed that the student could have faulty oral reading and still achieve comprehension. Later researcher (Powell, 1971) also found that errors in oral reading did not always affect comprehension. He studied the congruent validity of the Betts (1940) formula as far as the Word Recognition area was concerned, presuming the Comprehension Score of 70-75%. Powell (1969) argued that in spite of the fact that examiners used the same formula, discrepancies still arose. The problem was the procedural differences of examiners. Some examiners allowed students to read passages silently before reading them aloud, some had the students read only orally, others counted repetitions as errors.

Powell (1969) went further to say that if the Comprehension Level of 70-75% remained, the student could tolerate whatever Word Recognition Errors that accompanied that performance level. He found that younger children could tolerate more Word Recognition Errors than older children and still maintain the same level of comprehension. This would be true especially for non-English speakers; they could make numerous Word Recognition Errors--especially mispronunciations--and come out with a low Word Recognition Score; yet their comprehension could remain the same throughout different levels.

Powell (1971) further studied oral reading by analyzing five tests: *Spache Diagnostic Reading Scales*, *Durrell Analysis of Reading Difficulty*, the *Gilmore Oral Reading Test*, the *Gray Oral Reading Test*, and the *Gates McKillop Reading Diagnostic Test*. Two IRI's were also studied: McCracken's "Standard Reading Inventory" and Silvaroli's "Classroom Reading Inventory." Each test was analyzed for: (a) the number of words to be read in each passage (a count was made); (b) the maximum number of errors allowable for that passage to be regarded within limits of acceptable reading behavior from a given norm; and (c) Word Recognition, which was computed for each passage by dividing the number of words to be used by the number of errors allowable. The results were as follows: (a) all Word Recognition error ratios increased in error latitude as the difficulty of the materials increased and the age-grade of the sample increased; (b) one error in every 20 words (95% Word Recognition Accuracy) for determining Instructional Level is hardly justified because in an earlier study, Powell (1969) had discovered that some students could tolerate more oral reading errors than others and still comprehend.

More recently, researchers (Goodman and Burke, 1972) from a linguistic background, have questioned the rating or oral reading errors on IRI's. They have considered examining oral reading miscues on the basis of semantic, lexical, or syntactical errors. They present a more general schemata of reading miscue errors, which are based on dialect, intonation, graphic similarity, sound similarity, grammatical function, correction,

grammatical acceptability, semantic acceptability, and meaning change. Although this model of scoring oral reading has been criticized, it would seem to be more acceptable in purposefully identifying oral reading errors, especially for bilingual students; but for the purpose of this study, Goodman's Miscue Analysis (1972) was not used.

Since the concern of the researcher was to define reading in terms of comprehension, oral reading did not seem appropriate for ascertaining the bilingual student's comprehension level. For the purpose of this study, silent reading was deemed appropriate because the student would not be punished for pronunciation and accent. The student would be comfortable and would not feel embarrassed as in the case of oral reading, which many researchers (Otto *et al.*, 1973; Lowell, 1969; Roswell and Natchez, 1964) have viewed as a trying task for some students. Students usually concentrate intensively on the decoding task and neglect comprehension. As a result, it would then be unfair to estimate their comprehension based on oral reading.

In administering the IRI, it was decided to allow students to ready silently without a time limit per passage, although a time limit of the testing session was established. It was hoped that through a general time limit, the examiner could determine which students needed training for speed. It was also hoped that administration of the silent reading IRI would determine: (a) which students needed work in grasping the meaning of written materials (Bolenius, 1919); (b) the comprehension problems students were encountering; (c) the strengths

and weaknesses in the student's ability to follow directions, to read for understanding, and to use context (Dechant, 1970; Wilson, 1972); (d) specific skills in finding the main idea, character analysis, sequencing of details, making inferences, drawing conclusions, cause and effect, and higher-level questions asking the student to apply concepts to himself/herself (Valmont, 1972).

Asking the bilingual student, or any student for that matter, to apply concepts to his/her own experience helps make reading a more relevant experience. In fact, Knowles (1975) states that as a person grows, the concept of himself/herself moves from dependency to self-directedness. This self-direction is accompanied by a wealth of experience that can be a rich resource for learning and a readiness to learn, particularly those things that will help the student directly in daily problems.

The comprehension questions in the stories ask the student to relate original themes to his/her own life experience. This should not only facilitate the testing of comprehension, but should also provide the student with a positive attitude toward reading--the ultimate goal of reading.

Another facet of IRI's that has been controversial is the Word Recognition and Comprehension Scoring criteria to be used in determining the three functional reading levels: (a) Independent Level, (b) Instructional Level, and (c) Frustration Level. In studies since the 1940's, various criteria have been proposed for both word recognition and comprehension.

Early research in creating criteria for scoring IRI's was initiated by Killgallon and Betts (cited in Walter, 1974). Since Killgallon had worked closely with Betts at the Pennsylvania University Reading Clinic, the scoring criteria developed became known as the "Betts-Killgallon Criteria." The exact research base for the development of the criteria is not very well known, and for this reason they have been criticized. However, the criteria has generally remained applicable for IRI's.

Beldin (1970) suggested that the criteria was not arbitrarily established but was derived from a requirement of 50% comprehension for the understanding of material (Bolenius, 1919), and that for the student to read with meaning, he/she should not have more than one difficult word in 20 continuous words (Durrell, 1956). The Betts-Killgallon Criteria, which has been applied in most texts discussing IRI's, provides a Basal Reading Level of 90%, a Probable Instructional Level derived by a minimum comprehension score of 50%, a Probable Reading Capacity Level of 75%, and a Probable Frustration Reading Level attained by a comprehension score of 20% or lower.

Cooper (cited in Beldin, 1970) conducted a study with students of grades 1-6 to determine the relationship between the relevance of using symptoms of reading problems, demonstrated with certain materials, as a basis for predicting suitability of reading materials for the purpose of instruction; and to develop criteria that could be used to estimate the level of reading materials. Cooper used standardized tests and an IRI at the beginning. At the end of his six-month experimental

period he found that: (a) students who had made the most progress were those who had two or less errors in 100 running words of reading matter; (b) this group of students had been placed in appropriate instructional materials; and (c) primary students who made seven or more errors in 100 running words, and at the intermediate level, students who made 11 or more errors in 100 running words were placed in materials that were not suitable for their instruction. Therefore, Cooper set Comprehension Level Scores at 70% for primary and 60% for intermediate level students.

McCracken (1967) provided detailed criteria for counting Word Recognition errors, and he redefined the criterion levels as: (a) Independent Level, when the score of every passage has been rated as Independent Level; (b) Instructional Level, when half of the scores fall below the questionable half of the Instructional Level; and (c) Frustration Level, when the scores of one passage is rated at the Frustration Level. The percentages proposed for Word Recognition are: (a) Independent Level: 99-100%, (b) Instructional Level: 95-98%, and (c) Frustration Level: 94% or less. For Comprehension, the scores are: (a) Independent Level: 90-100%, (b) Instructional Level: 51-89%, and (c) Frustration Level: 50% or less.

Kender (1968) compared the results of three instructional reading levels designated by three IRI's using three different criterion measures. The results of the three IRI's revealed significant differences among the means of a group of eighth-grade students. It seemed that a discrepancy among test re-

sults would have been obvious is tests, calculation of errors, scores, and criteria were all compared. However, among the IRI's in this study, too many factors were being considred to make valid conclusions about them.

Powell (1969) disputed the Betts-Killgallon Criteria for scoring IRI's and said it was too high. He maintained through his own studies that pupils in grades one and two could tolerate an 85% Word Recognition Score and still comprehend 70% of the material; and students in grades three through six could attain a 91-94% score on Word Recognition, and comprehend 70% of the material at the same time.

Ekwall *et al.* (1973) conducted a study using a polygraph to validate the criteria for scoring IRI's. Their scope was to see if one set of criteria was more applicable to students of various mental ages, sexes, ethnic backgrounds, reading levels, and personality. To execute the study, they sampled 150 students of the third, fourth, and fifth grades. The students were administered an IRI individually, and they were monitored and taped by the polygraph to note their behavior during oral reading and responding to comprehension questions. Results were analyzed, and it was found that: (a) there was a significant difference in polygraph-measured frustration reading levels for third, fourth, and fifth graders; (b) there was no significant difference in polygraph-measured frustration reading levels between boys and girls; (c) there was no significant difference in polygraph-measured frustration reading levels among ethnic groups; (d) a comprehension criterion of

50% was adequate; (e) good readers become frustrated with fewer oral reading and comprehension errors than poor readers. In conclusion, Ekwall *et al.* did support Bett's Criteria for determining the Frustration Level of their students.

In light of all the studies on scoring IRI's, for the purpose of this study the researcher has set the criteria of: (a) 90-100% for the Independent Level; (b) 75-89% for the Instructional Level; and (c) 0-74% for the Frustration Level in determining the comprehension levels of bilingual students. Motta *et al.* (1974) employed a similar criteria for their non-English speakers.

Another problematic area to be addressed is who is qualified to administer IRI's? The philosophy of IRI's contend that the IRI can best be utilized by the classroom teacher. However, research has indicated that classroom teachers need training in construction, administration, and scoring of IRI's in order to effectively utilize them in the classroom. Perhaps teachers who are intimidated by their inadequate knowledge in readability, reading skills, taxonomies, methods of assessment, as well as prescription of remediation for students, are hesitant about employing IRI's in the classroom to: (a) match reading materials to their students' levels, (b) diagnose students' reading strengths and weaknesses, and (c) group students for reading instruction.

The following studies reported data on the teacher as the crucial factor in diagnosis, and the effectiveness of teachers before and after receiving training in administering IRI's. Sipay (cited in Durkin, 1970) reported that one consistent

finding from research is that the teacher is the crucial factor when student's achievement in reading is concerned. Durkin (1970) contended that research findings do not place enough emphasis on the teacher factor, and do not closely speculate what makes a successful or unsuccessful teacher. She questioned diagnosing reading levels until the teacher's knowledge of reading skills has been affirmed because the success or failure of diagnosis and interpretation of results for grouping and instructing students is dependent upon the teacher's knowledge. Farr (1970) claimed that the validity of the IRI is highly dependent upon the ability of its instructor and administrator.

Negative results have been reported with respect to teachers' ability to diagnose reading abilities before formal training. Emans (cited in Lowell, 1969) conducted a study rating 20 teachers in a graduate remedial reading course on their ability to distinguish reading skills needed by their pupils, whom they tutored one hour per day for five weeks. He found that the teachers had preconceived notions as to what skills the students needed, and were not perceptive nor accurate in their diagnosis of students' abilities.

Mills (cited in Lowell, 1969) reported that experienced classroom teachers were not aware of the frustration reading levels among their students when they were asked to estimate the frustration levels by using IRI techniques.

In order to determine whether teachers who had limited training in administering and scoring IRI's could assign appropriate instructional levels to students as accurately as clini-

cians, Blynn (1970) designed a correlational study. She had alternate forms of the "Standard Reading Inventory," Forms A and B administered to 30 students by three pairs of teachers and clinicians. Her findings indicated that differences in reader levels between pairs of examiners existed in 24 out of 30 cases. This supported that fact that classroom teachers need training in administering and scoring IRI's for assigning appropriate instructional levels to their students in grouping them for reading or meeting the students' instructional needs.

On the other hand, the following studies presented positive results in that teachers who were trained were able to effectively utilize IRI's. Kelly (1969) conducted the Berea, Ohio In-service Educational Experiment to investigate the importance of training teachers for the purpose of discovering the relative effectiveness of an adopted model of the IRI Instructional Process as a means of helping teachers become more cognizant of instructional reading levels in the classroom, and whether the time of the school year when the teacher participated in the in-service made a difference in the teacher's awareness of the instructional reading levels of students in the classroom. The results were:

1. In terms of evaluative basal reading materials, teachers who participated in simulation type in-service programs early in the school year were more aware of the instructional reading levels of students in the classroom.
2. There was a considerable difference between the teachers who had participated in the simulation in-service pro-

gram later in the year and teachers who did not participate at all.

3. Teachers who participated in the in-service program early in the year were significantly more aware of instructional reading levels than teachers who did not participate.
4. Primary school teachers were more aware of instructional reading levels in classrooms than intermediate school teachers (Kelly, 1969, p. 7).

The reason was that elementary teachers, in contrast to intermediate teachers, have to teach reading and are aware of the process. Nonetheless, the implication here is obvious: teachers do need training in preparing, administering, and interpreting IRI's; but most important, they need some background in the teaching of reading.

Ladd (cited in Lowell, 1969) re-affirmed that teachers need training in reading instruction through a study that found that after 30 hours of intensive training, teachers were able to evaluate the reading performance of students accurately.

For the purpose of this study, to ensure that classroom teachers of bilingual students effectively used the GIRI, a teacher-training workshop was conducted. The focus of the workshop was the reading process for bilinguals as well as administering, scoring IRI's, and prescribing remediation.

While literature dealing with the relationship of the IRI and *Cloze* tests used with English speakers is limited, literature dealing with the same relationship with non-native speakers

is practically non-existent. Research has reported that *Cloze* tests can be used as substitutes or partial substitutes for the IRI (Burmeister, 1974). This concept was elaborated by Bormuth (1967) who also found that the comprehension scores of IRI's and *Cloze* tests correlated significantly for English speakers. Similar research for non-English speakers was not found. However, it was reported that the IRI was effective in determining reading levels of non-native speakers (Motta *et al.*, 1974) and the *Cloze Test* was an effective measure for determining the comprehension levels of non-native speakers (Jongsma, 1971; Aitken, 1977; Oller and Conrad, 1971; Oller, 1972; Stubbs and Tucker, 1974). This led the researcher to hypothesize that perhaps the comprehension scores of IRI's and *Cloze* tests would correlate for non-native speakers. This hypothesis was based on Bormuth's study (1967), which will be discussed along with other studies reporting on the relationship of IRI and *Cloze* tests.

Bormuth (1967), the reading expert widely identified with *Cloze* procedures, conducted three studies based on Taylor's work (1953). Bormuth's first study examined the relationship of comprehension scores for multiple-choice and *Cloze* tests, constructed for each of his nine passages. The *Cloze Test* was given and within three days the multiple-choice tests were administered to 100 fourth and fifth graders. The reason for this time lapse is that the same passages were used for the multiple choice and *Cloze* tests. When scores were collected and compiled, a significant correlation was found between scores on the *Cloze* and the multiple-choice tests.

For the purpose of this study, the Bormuth study was replicated. An original IRI consisting of eight passages was prepared. From the GIRI stories, a *Cloze* test and a multiple-choice test were prepared for each of the eight passages.

In another study, Bormuth (1969) attempted to examine factors in the validity of the *Cloze* test as a measure of reading comprehension. A series of passages were leveled according to the Dale-Chall's Readability Formula, and a *Cloze* test and multiple-choice IRI were constructed for each passage. The tests were administered to 150 fourth, fifth, and sixth graders. Only the exact words omitted were accepted as correct on the *Cloze Test*. Conclusive results indicated that there was a high correlation between scores on the multiple choice and the *Cloze Test*.

Bormuth (1968) conducted another study to determine a set of criterion scores for *Cloze* readability tests that would be comparable to the criterion scores used with oral reading tests to determine the readability of passages, and to further substantiate that *Cloze* tests are measures of comprehension. He tested 120 pupils of grades four, five, and six, using four forms of the *Gray Oral Reading Test* (Robinson and Gray, 1963), each containing passages leveled from primer to high school level. He also administered two *Cloze* tests. Each student was administered two *Cloze* tests and an oral reading test, which tested word recognition and comprehension. The *Cloze* test scores correlated highly with the comprehension scores of the oral test. Comparable criteria were determined with *Cloze*

scores of 44% and 57% corresponding with comprehension criterion scores of 75-90%; and Cloze scores of 33-54% corresponding with word recognition scores of 95-98%.

Other evidence also indicates that there is a correlation between the *Cloze Test* and the IRI (Wiechelman, 1971). Wiechelman conducted a study comparing the functional reading levels identified with a *Cloze Test* and the functional reading levels identified with an IRI for 71 eighth-grade students, of which 13 were Spanish-surnamed. The conclusions indicated that there was a positive relationship between the functional reading levels identified by the *Cloze Test* and by the IRI. It was found that the mean functional reading levels through the use of the *Cloze Test* for these students did estimate their mean functional reading levels as reported by the IRI. Also, the mean instructional reading levels from the *Cloze Test* for Spanish-surnamed students did approximate their mean instructional reading levels on the IRI. Finally, when the instructional levels of the IRI and *Cloze* tests were compared with the *Durrell Listening Reading Test*, results indicated that the IRI and *Cloze* tests were more accurate in identifying the reading levels.

To demonstrate the effectiveness of the *Cloze Test* with non-native speakers, Oller and Conrad (1971) constructed a *Cloze Test* and administered it to 102 foreign students entering UCLA. Only the exact words omitted were counted as correct in the scoring. The students were also administered the UCLA *English as a Second Language Placement Examination* (ESLPE) as a basis of comparison with the *Cloze Test*. The researchers found the highest corre-

lation between the *Cloze Test* and the *Dictation* sub-test (which measured listening comprehension), and the next highest correlation between the *Cloze Test* and reading. These correlations led the researchers to recommend the *Cloze Test* as a good method for measuring language proficiency and comprehension levels, which can be used in placing non-native English speakers in English and reading classes.

Research on the IRI's has also focused on the relationship between scores on IRI's and on standardized reading tests. It has been generally discovered that standardized reading tests tend to overestimate the students' functional reading levels, and IRI's are more accurate measures for placing students at their appropriate reading levels.

Betts (1940) attempted to study the accuracy of standardized measures as compared to informal procedures for assessing reading grade placement. He administered five silent reading tests: the *Gates Reading Survey*, the *Stanford Achievement Test*, the *Durrell-Sullivan Reading Achievement Test*, the *Sangren-Woody Reading Test*, and the *Iowa Silent Reading Test--Advanced*. These were used to test fifth graders, and their scores were compared with the scores on the author's constructed IRI. Generally, Betts found the IRI to be a more accurate measure of reading levels.

Sipay (1964) studied the levels of reading achievement as measured by standardized reading tests and those levels determined by an IRI. He administered three standardized tests and two parallel forms of an IRI, and concluded that the three stan-

dardized tests tended to overestimate the instructional level of students one or more grade levels. Only two of the standardized tests, the *Metropolitan Reading Test* and the *Gates Reading Survey*, appeared to indicate the instructional reading levels when a more stringent criterion was applied.

McCracken (1962) has recommended the use of IRI's rather than standardized group reading tests for the purpose of obtaining instructional reading levels. He compared the performance of 56 sixth-grade pupils on the *Iowa Every-Pupil Test of Basic Skills, Test A: Silent Reading Comprehension* to the reading comprehension and vocabulary scores on an IRI, which included oral and silent reading. The three levels of performance on the IRI were the Immediate Instructional Reading Level, the Maximum Instructional Reading Level, and the Word Recognition Level. Results indicated that the average difference betweeen the *Iowa Reading Comprehension* grade levels and those estimated by the IRI was 2.3 years. The difference for the Vocabulary Grade Score was one year, and was higher for the *Iowa Test*. McCracken (1962) concluded that the use of the standardized test scores to determine the level of instruction would place 63% of the students at a Frustration Reading Level. He suggested that the instructional level be two grades below the standardized test scores. McCracken's suggestion has validity only for the *Iowa Every-Pupil Test of Basic Skills*, which he used in his study, and the reading materials that formed the basis of his IRI.

Williams (cited in McCracken, 1962) compared the performance of fourth, fifth, and sixth graders on an IRI based on their classroom basal readers with their scores on the *California Reading Test*, the *Gates Reading Survey*, and the *Metropolitan Achievement Reading Tests: Reading*. When an IRI, containing selections from basal readers which the students were familiar with, was used, the standardized tests were found to place students relatively near their instructional levels.

Brown (1963) designed a study to determine if a difference existed in the instructional reading levels as indicated by an IRI and the grade placement reading scores of five standardized reading tests. She administered the *California Reading Test*, the *Metropolitan Reading Test*, the *Stanford Reading Test*, the *Iowa Test of Basic Skills*, the *Gates Reading Test*, and an IRI to 192 elementary school children. After all the tests were administered, a second IRI was given to 49 of the students for the purpose of establishing the reliability of the IRI. When teachers were asked to estimate students' reading levels, it was found that their prediction correlated only with the IRI. Finally, results on the two IRI's correlated highly; and no significant difference was found between pupils' scores on the two forms.

Sipay (1964) attempted to obtain evidence on the extent to which the level of reading achievement, as measured by standardized reading achievement test scores, differed from functional reading levels, as estimated by an author-constructed IRI. He administered the *Metropolitan Achievement Test: Reading*,

the *Gates Reading Survey*, and the *California Reading Test* to 202 subjects from 8 fourth-grade classes. The students were given individually administered IRI's, which were based on selections from the Scott Foresman Reading Series. The criteria for determining the functional reading levels were as follows: for the Instructional Level: Cooper Criteria--96% with Word Pronunciation 96-99%, and Comprehension--minimum 60%; for the Betts Criteria--90% with Word Pronunciation 90-95%; and Comprehension--minimum 60%; for the Frustration Level: Word Pronunciation less than 90%, and Comprehension--minimum of 50% or less. The statistical analysis of the test scores indicated that:

1. In estimating with the Cooper Criteria, the Instructional Level indicated by all three standardized tests tended to overestimate the instructional level by approximately one or more grades.
2. When the Betts Criteria were used, the mean score of the *Metropolitan Test* was .11 grade levels higher, while the *Gates Reading Survey* overestimated the Betts Criteria Instructional Level by .29 of the grade level, and the mean of the *California Reading Test* was 1.02 higher than that of the Betts Criteria Instructional Level.
3. The standardized tests, when compared with Frustration Level Criteria, were significantly lower in the case of the *Metropolitan* and *Gates* tests.
4. A comparison of the means of the Frustration Level and the *California Test* revealed that the *California Reading*

Survey underestimated the Frustration Level by .24 of the grade level. These differences were significant at the .05 level.

In conclusion, Sipay (1964) stated that these findings suggest that is is impossible to generalize as to whether standardized reading achievement test scores tend to indicate the Instructional or Frustration levels; rather, it appears that in making such a judgment one must consider the standardized reading test used and the criteria employed to estimate the functional reading levels.

In a study of the relationship of pupils' scores from IRI's and standardized tests, Glaser (cited in McCracken, 1962) compared the functional reading levels of retarded seventh grade and advanced third-grade students to their scores on the *Gates Reading Survey*. All the students in both groups had scored between 5.0 and 5.9 on the *Gates Reading Survey*. The result was that standardized reading tests tended to overestimate reading levels. The findings of this survey were:

1. The Instructional Levels of the advanced and retarded readers were consistently lower than the levels of their standardized reading test scores with a slightly larger spread evident for retarded readers.
2. Sixteen (52%) of the retarded seventh-grade readers reached Frustration Level in passages of fifth grade difficulty. Seventeen (50%) of the third-grade pupils met the criteria for Frustration at this level.

3. The Instructional Levels were consistently below the standardized reading test scores for the two groups.
4. Providing reading instruction and materials for students on the basis of standardized reading test scores could hinder their progress and possibly affect their attitude toward reading.

Leibert (cited in McCracken, 1962) compared the scores of IRI's and the *Gates Advanced Primary Reading Test* for second graders. Leibert reported differences in grade placement for the two measures but suggested that these differences may be due to the wider range of skills included in a group standardized test, while reading as measured by the IRI is more narrowly defined.

Patty (1965) contrasted scores on the *Gilmore Oral Reading Test* and the *Gray Oral Reading Test* with the IRI performance. Patty found that it was impossible to generalize as to whether standardized oral reading tests indicate the functional reading levels of children as accurately as IRI's do. Because of the economy of administration and the usefulness of the information they provided, the *Gray Oral Reading Test* and the IRI were deemed the most desirable instruments for determining functional reading levels.

Brown (1963) came to a similar conclusion in a study using the following silent reading tests: the *California Reading Test*, the *Metropolitan Achievement Test: Reading*, the *Stanford Achievement Test: Reading*, the *Iowa Every-Pupil Test of Basic Skills*, and the *Gates Reading Survey*. Brown found no consistent relation-

ship between performance on these tests and on informal inventories. The Brown and Patty studies are not directly comparable; Brown used standardized silent reading tests, while Patty used standardized oral reading tests.

Botel (1969) confirmed that standardized reading tests have not been accurate in indicating at what level the student should be reading. As a result, 10 to 15 million students in United States schools are reading books that are beyond their instructional levels. Botel (1961) explained that he and other reading authorities agreed that the reading of children forced to read books that are too difficult for them is affected negatively.

Farr (1969) contended that standardized test scores were almost useless for the diagnosis of students' instructional reading levels in the classroom. More specifically, Burmeister (1974) asserted that standardized silent reading tests have little diagnostic value because only a limited number of any one type of items were included in any one form of these tests; hence, they provided the teacher only with a grade or pencentile score that ranks students according to norms for a given population. The scores obtained from these silent reading tests by students above the primary grades tended to overestimate reading levels and place students at their frustration levels. A teacher has to judiciously subtract a year from each student's total score to estimate the student's instructional levels and two years to estimate the student's independent level.

Smith (1970) compared the results for fourth-, fifth-, and sixth-grade subjects from the administration of the Vocabu-

lary and Comprehension sub-tests of the *Gates MacGinitie Reading Tests, Survey D*, The Vocabulary Reading and Paragraph Reading sub-tests of the *Durrell Listening and Reading Series, Intermediate Level* with those of an IRI. The findings demonstrated that for all three grade levels there were no statistically significant differences between the mean grade scores obtained from the standardized tests and the mean instructional reading levels obtained from the IRI. However, the *Gates MacGinitie* and *Durrell* tests placed more than one-half and one-third of the subjects, respectively, from one to five years above their actual instructional reading levels. Neither the sub-tests nor the total score from the two standardized tests was accurate in estimating the students' instructional reading levels for pupils of grades four, five, and six.

A study was conducted by Feldman *et al.* (1971) to find if results from different measures could be used for various purposes, such as placing the students in groups, or for selecting appropriate reading material according to their reading levels. They tested 96 children of grades one to three with two standardized measures: the *New York Growth in Reading*, the *Metropolitan Achievement Test*, and two non-standardized tests--the "Harris Sample Graded Word List," and "Graded Basal Readers Test" used in the students' reading program. The results indicated that the students' scores were higher on the standardized tests, which tended to overestimate their functional reading levels in the classroom; and the IRI's seemed better suited for measuring the instructional reading levels in the classroom.

The four tests appeared to measure different skills; the non-standardized measures tested specific sight vocabulary, while the standardized tests measured global reading skills.

Wade (1971) tested 77 eighth-grade students to compare their reading scores on the *Gates MacGinitie Reading Test, Survey E*, and the *Durrell Listening Series, Advanced Level, Form D-E*, and an IRI. The mean grade scores obtained from the sub-test and from the total test scores of the standardized reading tests tended to overestimate the instructional reading levels for these students by two and a half years.

O'Brien (1973) also contended that standardized tests measured only global reading levels, while IRI's provided the basis for estimating Independent, Instructional, and Frustration Reading Levels. Scores on standardized tests gave the teacher information about a group of students in relation to age-grade norms, but the scores did not delineate the nature of the student's problems in reading. On the other hand, scores on the IRI provided the teacher with specific information on the students' ability to attack unknown words, their ability to read with comprehension, and the levels at which the students were capable of performing in these two areas.

From these studies it can be surmised that standardized tests are not the best measures for assessing students' functional reading levels. However, the problem is further compounded if the reading levels of bilingual students are assessed using standardized reading tests. First of all, norms may have been established with native English speakers. Secondly, items

on these tests may be culturally biased in that the tests presuppose certain visual and auditory skills for which the students may not have received adequate training. Thirdly, bilingual students are made to feel deficient in tests of unfamiliar lexical items and syntactical structures.

Motta *et al.* (1974) recommended the use of IRI's with ESL and bilingual students because they claimed that standardized tests did not take into account such factors as socio-economic status, motivation, culture, nor the psycholinguistic experience of these students. In their study, they prepared two sets of graded passages and corresponding comprehension questions for these paragraphs. One set of passages were administered individually and the other as a group test to non-English-speaking adults. Both tests successfully indicated that students' functional reading levels. They advised that to further estimate the student's literacy in his/her own language, an IRI in the student's native language should also be administered.

PLAN AND PROCEDURE OF THE STUDY

Sample Population

Fifty bilingual students of predominantly English and Hispanic-language backgrounds (with the exception of two Italian children and one Lebanese child) were randomly selected for this study. The subjects lived in an area that included people of low and middle socio-economic (SES) backgrounds. They attended a city public school, which provided them with a bilingual program of Spanish and English and an

ESL program. All students were instructed in the content area subjects in Spanish until they gained language proficiency in English. As soon as the students were capable of functioning at grade level in English, they were removed from the bilingual and mainstreamed into the English curriculum.

Instruments

1. A Group Informal Reading Inventory (GIRI) was constructed as a set of eight passages based on original themes. The passages ranged in length from 30-200 words. The *Fry* and *Dale-Chall Readabilities* were used to level the passages from primer to eighth grade. Ten multiple-choice comprehension questions were constructed for each passage based on the Barrett (cited in Lapp and Ramsey, 1976) and Krathwohl (1964) Taxonomies.

   The Informal Reading Inventory (IRI) was administered as a group test. To score the test, a percentage was determined and compared to the following criteria in order to determine the comprehension reading levels: (a) Independent: 90-100%, (b) Instructional: 75-89%, and (c) Frustration: 0-74%.

2. *A Cloze Test* was constructed by deleting every seventh word. The first and last sentence of each passage were left intact. The test was administered as a group test and students were asked to fill in the blanks. In this test an answer was considered correct

as long as it was appropriate contextually. To score the *Cloze Test*, Bormuth's criteria (1967) was used: (a) Independent Level: 50-57%, (b) Instructional Level: 35-46%, and (c) Frustration Level: 19-31%.

3. The *Stanford Diagnostic Test* (1976): Red (grades 1-3), Green (grades 3-5), and Brown (grades 5-8) levels were used to examine literal and inferential reading comprehension.
4. A questionnaire entitled "Teacher Estimate of the Students' Reading Levels" was used. Bilingual students' names were listed and the teacher was asked to estimate the reading level (grade) at which each student was functioning. (This was considered the instructional reading level.)

Procedure

After the population was selected, arrangements were made with the bilingual coordinator of the school to conduct a teacher-training workshop and to administer the tests. The teacher-training workshop conducted for six bilingual teachers included: (a) a presentation and explanation of the construction and purpose of the tests to be administered; (b) a practice administration, scoring, and evaluation of the tests among the teachers; (c) a discussion and planning of the group testing with bilingual students; and (d) a completion of the "Teacher Estimate of the Students' Reading Levels" questionnaire. The arrangements were that the *Cloze Test* would be administered

initially; the IRI would follow within three days, as exemplified in the Bormuth study (1967); then, the *Stanford Diagnostic Test* (1976) would be administered.

Summary and Analysis of Data

Data concerning the demographics and results from the assessment measures administered to bilingual students are summarized in Table 1.

The scores from the *Stanford Diagnostic Test* (1976) are an average of the literal and inferential comprehension results for the red, green, and brown levels, respectively. For desciptive purposes, the students' report card grades on Reading in the Native Language (N.L.) and English as a Second Language (ESL) were also included.

Close inspection of Table 1 demonstrates that this random sample of bilingual students is functioning not only far below their developmental grade level but also below their assigned present grade level. Interestingly enough, this fact holds true even for students who have been exposed to English reading instruction up to five years. Clearly then, the original argument that there is a pressing need to assess these students' functioning reading levels as early as possible and to match instruction and materials to the students' reading needs is valid.

In noting the students' reading grades, those in the native language appear to be much lower than those in English as a Second Language. A possible reason for this is that students

TABLE 1

DEMOGRAPHICS AND RESULTS FROM ASSESSMENT MEASURES ADMINISTERED TO BILINGUAL STUDENTS

					Instructional Estimated Grade Level						
								Stanford		Previous Grades	
Subject	Age	Grade Student Should Be In	Present Grade	Years in U.S.	Teacher Estimate	Cloze	GIRI		Level	N.L.	ESL
1	14	9	8	5	8	5	7,8	6.6	*	B	A
2	13	8	7	2.5	3-5	1	5	5	✓	A	A
3	14	9	7	3	1-2	P	4	2.7	X	D	B
4	14	9	8	3	3-5	1	4	3.6	✓	-	A
5	14	9	7	5	3-5	5	4	3.7	✓	B	A
6	13	8	7	3	1-2	4	6	4.8	X	-	B+/A
7	15	10	8	5	3-5	4-5	5	4.5	*	C	B+/A-
8	14	9	8	3.5	3-5	3	4	4	✓	C	A
9	13	8	7	4	1-2	1	4	2.9	✓	-	B
10	15	10	8	2	1-2	2	7,8	4.5	*	A	A

Key: N.L. = Native Language--Reading
ESL = English as a Second Language
P = Primer Level
* = Stanford--Brown Level
X = Stanford--Green Level
✓ = Stanford--Red and Green Level

are heterogeneously grouped in the native language, and thus each student has to compete with students of all ability levels; whereas, for English as a Second Language, students are homogeneously grouped competing only with students who normally function at their own level.

Contradictory to the above, both native- and second-language teachers tend to underestimate the students' reading levels when compared with the results of reading tests. This factor could be related to the students' poor performance in reading--the self-fulfilling prophecy--teachers do not feel the students can do well, and as a result, students do not do well. Ironically, ESL teachers assign "A's" and "B's" to the students on the report cards. The question is "*Why*?" If the rationale is to help the students' self-concept, then this is acceptable. However, it cannot be overlooked that they may simply be complying to the educational system's social promotion policy, in spite of the fact that they are aware of the students' low reading levels. In this case, the students are deluded in believing that they are achieving and may not be prepared in terms of reading skills to compete with the world in and out of school.

IMPLICATIONS FOR TEACHING

Clearly, it is the case that when junior high to high school bilingual students are functioning five or more years below grade level, something has to be done about instruction. The danger is obvious if instruction in the native language

and English is not presented as a relevant experience; these students become frustrated and eventually drop out of school. Teachers must consciously deal with the students' reading needs and not gloss over these needs by issuing good grades. To deal with these needs, teachers must meet the students' (a) academic needs by providing high interest, low level materials and (b) the students' interests by providing materials and skills that will prepare them for whatever endeavor they seek in the outside world whether it be higher education, academic, or a working vocation.

The possible guidelines for meeting student needs are:

1. Language medium--A decision has to be made whether the student is to be instructed in two languages or whether it may be too late for the student who does not have native language competencies to get the skills in the native language and transfer these. ESL may be the solution in this case because if the student is frustrated, he/she will drop out of school altogether.

2. Assessment--Pretests and post tests (informal and formal measures) must be administered on a continuing basis so that instruction can constantly be adjusted to the students' reading and content area, strengths, and weaknesses.

3. Instruction--(a) *Time* to be allotted in the two languages; (b) *Technique* to be employed--whole group instruction, individualized. (The latter would be the most beneficial in placing part of the responsibility of instruction to the student. With the issuing of performance contracts, the student

would then be accountable for his/her success and failure.); (c) *Content* relevant, functional instruction. For example, learning a driver's manual, filling out job applications, or relevant information that the student needs to deal with in his/her English environment; (d) *Materials* commerical and teacher- or student-made, high interest, low reading level; (e) *Evaluation* of the student, instruction, program, and changes made accordingly.

REFERENCES

Aitken, Kenneth G. "Using Cloze Procedure as an Overall Language Proficiency Test," *TESOL Quarterly*, XI, No. 1 (March, 1977), 59-67.

Anderson, Theodore, and Mildred Boyer. *Bilingual Schooling in the United States.* Austin: Southwest Educational Laboratory, 1970.

Beldin, H. O. "Informal Reading Testing: Historical Review and Review of the Research," *Reading Difficulties: Diagnosis, Correlation, and Remediation.* Newark: International Reading Association, 1970.

Betts, Emmett A. *Foundations of Reading Instruction, With Empasis on Differentiated Guidance.* New York: American Book Co., 1957.

_____. "Reading Problems at the Intermediate-Grade Level," *Elementary School Journal*, XL, No. 10 (June, 1940), 737-746.

_____. *The Prevention and Correction of Reading Difficulties.* Evanston: Row Peterson & Co., 1936.

Bloom, Benjamin S., ed. *Taxonomy of Educational Objectives: The Classification of Educational Goals, by a Committee of College and University Examiners.* New York: Longmans, Green, 1956.

Blynn, Catherine Peters. "A Determination of the Practicality of Making the Use of an Informal Reading Inventory More Applicable to the Needs of the Classroom Teachers," *Dissertation Abstracts International.* Lehigh University. XXXIA, No. 10, Part 5, 1970, p. 5240-A.

Bolenius, Emma M. *Teacher's Manual of Silent and Oral Reading.* Boston: Houghton Mifflin, 1919.

Bormuth, John R. "Cloze Test Reliability: Criterion Reference Scores," *Journal of Educational Measurement*, V, No. 3 (Fall, 1968), 189-196.

_____. "Comparable Cloze and Multiple-Choice Comprehension Test Scores," *Journal of Reading*, X, No. 5 (February, 1967), 291-299.

_____. "Factor Validity of Cloze Tests as Measures of Reading Comprehension Ability," *Reading Research Quarterly*, IV, No. 3 (Spring, 1969), 358-365.

Botel, Morton. *The Botel Reading Inventory*. Chicago: Follett, 1961.

_____. "The Validity of Informal Reading Testing." Paper presented at the 14th Annual Convention of the International Reading Association, Kansas City, 1969.

Brown, Sandra Rose. "A Comparison of Five Widely Used Standardized Reading Tests and an Informal Reading Inventory for a Selected Group of Elementary School Children," *Dissertation Abstracts International*. University of Georgia. XXV, No. 1, Part 1, 1963, p. 996.

Brown v. Board of Education of Topeka, 347 U.S. 483 (1954).

Burmeister, Lou. *Reading Strategies for Secondary School Teachers*. Menlo Park: Addison Wesley Publishing Co., 1974.

Condon, Elaine. "The Cultural Context of Language Testing," *Papers in Language and Testing*, eds. Palmer, and Spolsky. Washington: Teachers of English to Speakers of Other Languages, 1975.

Dale, Edgar, and Jeanne S. Chall. "A Formula for Predicting Readability," *Educational Research Bulletin*, XXVII, No. 1 (January 21, 1948), 11-20, 28.

Dechant, Emerald V. *Improving the Teaching of Reading*. 2d ed. Englewood Cliffs: Prentice Hall, Inc., 1970.

Dunkeld, Colin George Mackinnon. "The Validity of the Informal Reading Inventory for the Designation of Instructional Reading Levels: A Study of the Relationships Between Children's Gains in Reading Achievement and the Difficulty of Instructional Materials," *Dissertation Abstracts International*. University of Illinois at Urbana-Champaign. XXIA, No. 12, Part 6, 1970, p. 6274-A.

Durkin, Dolores. *Teaching Them to Read.* Boston: Allyn and Bacon, 1970.

Durrell, Donald D. *Improving Reading Instruction.* New York: World Book Company, 1956.

Ekwall, Eldon E., Judy K. English Solis, and Enrique Solis, Jr. "Investigating Informal Reading Inventory Scoring Criteria," *Elementary English*, L, No. 2 (February, 1973), 271-274.

Farr, Robert C. *Measurement and Evaluation of Reading.* New York: Harcourt, Brace & World, Inc., 1970.

_____. *Reading: What Can be Measured?* Newark: International Reading Association, Research Fund, 1969.

Feldman, Shirley C., *et al. Analyzing Reading Growth of Disadvantaged Children Through Longitudinal Study of Several Reading Measures.* U.S., Educational Resources Information Center, ERIC Document, ED 047 900, February, 1971.

Flesch, Rudolf. "A New Readability Yarkstick," *Journal of Applied Psychology*, XXXII, No. 3 (June, 1948), 221-233.

Fry, Edward. "A Readability that Saves Time," *Journal of Reading*, XI, No. 7 (April, 1968), 513-516.

Goodman, Yetta, and Carolyn Burke. *Reading Miscue Inventory.* New York: The MacMillan Co., 1972.

Harris, Albert J. *How to Increase Reading Ability: A Guide to Developmental and Remedial Methods.* 4th ed. rev. New York: David McKay Co., Inc., 1961.

Herbert, W. "Bilingual Education Missing the Mark?" *APA Monitor*, VIII (1977), 8-9.

Johnson, Marjorie S., and Roy A. Kress. *Informal Reading Inventory.* Newark: International Reading Association, 1965.

Jongsma, Eugene. *The Cloze Procedure: A Teaching Technique.* Newark: International Reading Association, 1971.

Karlsen, Bjorn, Richard Natten, and Eric F. Gardner. *Stanford Diagnostic Reading Test.* New York: Harcourt Brace Jovanovich, Inc., 1976.

Kelly, Dean. *Using an Informal Reading Inventory to Place Children in Instructional Materials: The Berea, Ohio, In-service Education Experiment.* U.S., Educational Resources Information Center, ERIC Document ED 032 198, May, 1969.

Kender, Joseph P. "How Useful are Informal Reading Tests?" *Journal of Reading*, XI, No. 5 (February, 1968), 337-342.

Knowles, Malcom Shepherd. *Self-Directed Learning: A Guide for Learners and Teachers.* New York: Association Press, 1975.

Kobrick, Jeffrey W. "The Compelling Case for Bilingual Education," *Saturday Review*, LV (April 29, 1972), 54-58.

Krathwohl, David R., Benjamin S. Bloom, and B. B. Masia. *Taxonomy of Educational Objectives, II, Affective Domain.* New York: Daniel McKay Co., 1964.

Lapp, and Ramsey. *Reading Theory and Practice.* New York: MacMillan Co., 1976.

Lau v. Nichols, 414 U.S. 563, 566 (1974).

Lindvall, C. M. *Measuring Pupil Achievement and Aptitude.* New York: Harcourt, Brace & World, Inc., 1967.

Lowell, Robert, E. *Problems in Identifying Reading Levels with Informal Reading Inventories.* U.S., Educational Resources Information Center, ERIC Document ED 032 199, May, 1969.

McCracken, Robert A. "Standardized Reading Tests and Informal Reading Inventories," *Education*, LXXXII, No. 6 (February, 1962), 366-369.

_____. "The Development and Validation of the Standard Reading Inventory for the Individual Appraisal of Reading Performance in Grades 1 Through 6." Ann Arbor: University Microfilms, 1964.

_____. "The Informal Reading Inventory as a Means of Improving Instruction," *Perspectives in Reading, No. 8, The Evaluation of Children's Reading Achievement*. Newark: International Reading Association, 1967.

_____, and Neill Mullen. "The Validity of Certain Measures in an IRI," *Reading Difficulties: Diagnosis, Correction, and Remediation*. Newark: International Reading Association, 1970, pp. 104-110.

Morishima, James, and Donald Mizokowa. "Testing and Evaluating Asian-American/Pacific Islander Bilingual Students." Tacoma: Bilingual Education Technical Assistance Center, 1977, pp. 1-4. (Mimeographed.)

Motta, Janice *et al.* *Reading Evaluation for Adult Non-English Students*. U.S., Educational Resources Information Center, ERIC Document ED 098 513, 1974.

O'Brien, Carmen. *Teaching the Language-Different Children to Read*. Columbus: Charles E. Merrill Publishing Co., 1973.

O'Brien, Patti. *An Informal Reading Inventory--A New Approach*. U.S., Educational Resources Information Center, ERIC Document ED 046 672, May, 1970.

Office of Civil Rights. "Task Force Findings Specifying Remedies Available for Eliminating Past Educational Practices Ruled Unlawful Under *Lau* v. *Nichols*, 1975. (Mimeographed.)

Oller, John W., Jr. "Scoring Methods and Difficulty Levels for Cloze Tests of Proficiency in English as a Second Language," *Modern Language Journal*, LVI, No. 3 (March, 1972), 151.

_____, and Christine A. Conrad. "The Cloze Technique and ESL Proficiency," *Language Learning*, XXI, No. 2 (December, 1971), 183-195.

Otto, Wayne, Richard A. McMenemy, and Richard J. Smith. *Corrective and Remedial Teaching*. 2d ed. Boston: Houghton Mifflin, Co., 1973.

Patty, Delbert Lee. "A Comparison of Standardized Oral Reading Test Scores and Informal Reading Inventory Scores," *Dissertation Abstracts International*. Ball State University, XXVI, No. 9, Part 5, 1965, p. 5302-5303.

Pikulski, John. "A Critical Review: Informal Reading Inventories," *The Reading Teacher*, XXVIII, No. 2 (November, 1974), 141-151.

Powell, William. "Reappraising the Criteria for Interpreting Informal Inventories," *Reading Evaluation, 1968 Proceedings*, ed. Dorothy De Boer. Vol. XIII, Part 4. Newark: International Reading Association, 1969, pp. 100-109.

_____. "Validity of the IRI Reading Levels," *Elementary English*, ILVIII (1971), 637-642.

Robinson, Helen M., and William S. Gray, eds. *Gray Oral Reading Test*. Indianapolis: Bobbs-Merrill Co., Inc., 1963.

Roswell, Florence G., and Gladys Natchez. *Reading Disability: Diagnosis and Treatment*. New York: Basic Books, Inc., 1964.

Ruddell, Robert B. "The Effect of Similarity of Oral and Written Patterns of Language Structure on Reading Comprehension," *Elementary English*, XLII, No. 4 (April, 1965), 403-410.

Silvaroli, Nicholas J. *Classroom Reading Inventory*. Dubuque: William C. Brown Co., 1965.

Sipay, Edward R. "A Comparison of Standardized Reading Scores and Functional Reading Levels," *The Reading Teacher*, XVII, No. 4 (January, 1964), 265-268.

Smith, Charles LeRoy. "A Comparison of Selected Standardized Reading Test Scores and Informal Reading Inventory Results at Intermediate Grade Levels," *Dissertation Abstracts International*. University of Northern Colorado. XXXI, No. 8, Part 4, 1970, p. 4046-A to 4047-A.

Smith, Nila Banton. "What Have We Accomplished in Reading?--A Review of the Past Fifty Years," *Elementary English*, XXXVIII, No. 3 (March, 1961), 141-150.

Spache, George A. "A New Readability Formula for Primary-Grade Reading Materials," *Elementary School Journal*, LIII, No. 7 (March, 1953), 410-413.

Stubbs, Joseph B., and G. Richard Tucker. "The Cloze Test as a Measure of English Proficiency," *Modern Language Journal*, LVIII, No's 5-6 (September-October, 1974), 239.

Taylor, Wilson L. "'Cloze Procedure': A New Tool for Measuring Readability," *Journalism Quarterly*, XXX (Fall, 1953), 415-433.

Thonis, Eleanor. *Teaching Reading to Non-English Speakers*. New York: Collier MacMillan International, Inc., 1976.

Tucker, Elizabeth S. *Instructional Level in Reading: Including a Critique of the Informal Reading Inventory*. U.S., Educational Resources Information Center, ERIC Document ED 122 255, 1975.

United States Commission on Civil Rights. *Ethnic Isolation of Mexican-Americans in Public Schools in the Southwest*. A Research Report on the Mexican-American Study Series, Report No. 1. Washington: Government Printing Office, 1971.

Valmont, William J. "Creating Questions for Informal Reading Inventories," *The Reading Teachers*, XXV, No. 6 (March, 1972), 509-512.

Vaughan, Joseph, Jr. "Interpreting Readability Assessment," *Journal of Reading*, XVIV (May, 1976), 635-639.

Wade, Vern Joseph. "A Comparison of Selected Standardized Reading Test Scores and Informal Reading Inventory Results at Eighth Grade Level," *Dissertation Abstracts International*. University of Northern Colorado. XXXIIA, No. 7, Part 4, 1971, p. 3853-A.

Walter, Richard B. *History and Development of the Informal Reading Inventory*. U.S., Educational Resources Information Center, ERIC Document ED 098 539, 1974.

The Way We Go to School: The Exclusion of Children in Boston. Boston: The Task Force on Children Out of School, 1970.

Wiechelman, Duane Sylvester. "A Comparison of Cloze Procedure Scores and Informal Reading Inventory Results for Estimating Functional Reading Levels for Students at Eighth Grade Level," *Dissertation Abstracts International*. University of Northern Colorado. XXXIIA, No. 7, Part 4, 1971, p. 3583-A to 3584-A.

Wilson, Robert M. *Diagnostic and Remedial Reading for Classroom and Clinic*. Columbus: Charles E. Merrill Publishing Co., 1972.

APPENDIX

An Illustrative Story from the Informal Reading Inventory (In Cloze and Multiple-Choice Versions)

STORY III

"Luciano, wake up! Have you forgotten what today is?" whispered mother. I opened my sleepy eyes. I _____ that today my uncle and his _____ would be arriving from Italy. I _____ as I thought of the food _____ fun I would have. Then I _____ to my mother. She smiled and _____ me. I then said, "Mother, can _____ take my cousins to the Feast _____ St. Anthony?" Mother wrinkled her forehead _____ answered, "You must not go alone. _____ go together."

I jumped out of _____ and dressed quickly. I wore the _____ suit that mother had bought for _____ special day. I ran out to _____ kitchen. Anna was crying, as she _____ did when Mother combed and braided _____ hair. Father and Grandfather sat at _____ table drinking expresso coffee as they _____ of the olden days in Italy. _____ liked to listen, because I had been there. Grandfather always said, "America _____ a rich country today, but Americans _____ not have the respect in families _____ we Italians have." Grandfather said, "A _____ is important. He knows the way _____ should be in the family." As _____ talked, Grandfather noticed me and said, "_____, Luciano, show your grandpa what a _____ grandson he has. You will be a fine man and a papa someday too!"

STORY III

"Luciano, wake up! Have you forgotten what today is?" whispered mother. I opened my sleepy eyes. I remembered that today my uncle and his family would be arriving from Italy. I smiled as I thought of the food and fun I would have. Then I turned to my mother. She smiled and kissed me. I then said, "Mother, can I take my cousins to the Feast of St. Anthony?" Mother wrinkled her forehead and answered, "You must not go alone. We'll go together."

I jumped out of bed and dressed quickly. I wore the new suit that mother had bought for this special day. I ran out to the kitchen. Anna was crying, as she always did when Mother combed and braided her hair. Father and Grandfather sat at the table drinking expresso coffee as they talked of the olden days in Italy. I liked to listen, because I had never been there. Grandfather always said, "America is a rich country today, but Americans do not have the respect in families that we Italians have." Grandfather said, "A papa is important. He knows the way things should be in the family." As he talked, Grandfather noticed me and said, "Come, Luciano, show your grandpa what a fine grandson he has. You will be a fine man and a papa someday too!"

STORY III--QUESTIONS

1. Who was coming to visit Luciano?
   a. his uncle and his family
   b. his father and grandfather
   c. his sister and mother
   d. his sister and her family

2. Where did Luciano want to take his cousins?
   a. to the block party
   b. to the neighborhood fiesta
   c. to the feast of St. Anthony
   d. to the amusement park

3. What did Luciano's father and grandfather drink?
   a. wine
   b. coca-cola
   c. tomato juice
   d. expresso coffee

4. Why did Luciano like to listen to Grandfather?
   a. He wanted to go to Italy.
   b. He had never been to Italy.
   c. He liked Italy better than America.
   d. He wanted to live in Italy.

5. What did Luciano do after he jumped out of bed?
   a. He went to the kitchen.
   b. He listened to his grandfather.
   c. He dressed.
   d. He talked to his mother.

6. Why do you think Anna was crying?
   a. Her grandfather spilled coffee on her.
   b. She fell off a tree.
   c. She did not want to eat.
   d. She was in pain.

7. How do you know that Grandfather missed Italy?
   a. He wanted to go back to Italy.
   b. He always talked about Italy.
   c. He had relatives in Italy.
   d. He was rich in Italy.

8. How do you know that Mother was concerned about Luciano?
   a. She did not want him to go to the feast by himself.
   b. She always bought him new suits.
   c. She did not want him to have a good time.
   d. She let him do whatever he pleased.

9. According to Grandfather, what is the difference between Italians and Americans?
   a. Americans are poorer than Italians.
   b. Italians drink more expresso than Americans.
   c. Italians have better houses than Americans.
   d. The people of the two countries have different family values.

10. How does Grandfather feel about Luciano?
   a. He is proud of him.
   b. He is unhappy with him.
   c. He dislikes him.
   d. He feels that Luciano likes to fool around too much.